

EP009/2053

2 3 0



REC'D 28 AUG 2000
WIPO PCT

EPO - DG 1

07. 07. 2000

(59)

**Prioritätsbescheinigung über die Einreichung
einer Patentanmeldung**

Aktenzeichen: 199 11 130.8
Anmeldetag: 12. März 1999
Anmelder/Inhaber: Jörg H a g e r , Bonn/DE
Bezeichnung: Verfahren zur Identifikation chromosomaler
Regionen und Genen
IPC: C 12 Q 1/68

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Anmeldung.

München, den 29. Juni 2000
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Hoiß

**PRIORITY
DOCUMENT**

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Zusammenfassung

5

10

Die vorliegende Erfindung betrifft ein Verfahren zum Nachweis von chromosomalen Fragmenten, welche eine Rolle in der Entwicklung komplexer Phänotypen, wie multifaktorieller Krankheiten spielen. Das Verfahren erlaubt den direkten Nachweis von DNA Fragmenten, welche in den Genomen zweier verwandter Individuen durch Abstammung identisch sind. Weiter erlaubt das erfindungsmäßige Verfahren die Selektion solcher DNA Fragmente, welche spezifisch für einen bestimmten Phänotyp sind. Darüberhinaus erlaubt das erfindungsmäßige Verfahren die Identifikation von Genen in Regionen, welche für einen bestimmten Phänotyp mitverantwortlich sind.

Stand der Technik

20

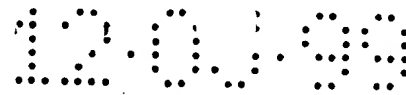
25

30

Eine der großen Herausforderungen der modernen Biologie und Medizin ist die Aufklärung der molekulargenetischen Grundlagen, welche bei der Entwicklung komplexer Phänotypen, wie sie multifaktorielle Krankheiten wie Diabetes, Adipositas, Asthma etc. darstellen, eine Rolle spielen. Die Grundlagen für die molekulargenetischen Untersuchungen dieser Krankheiten stammen weitestgehend aus dem Erfahrungsschatz zur Identifizierung seltener, monogener Erkrankungen, in denen ein klarer, kausaler Zusammenhang zwischen einer oder mehreren Mutationen in einem bestimmten Gen und einem Krankheitsphänotyp vorliegen. Dieser kausale Zusammenhang ist bei den weitverbreiteten Krankheiten, für die oben einige Beispiele angeführt wurden, nicht mehr gegeben. Bei solchen Krankheiten führen Veränderungen an vielen genetischen Loci zu einer Prädisposition für einen bestimmten Krankheitsphänotyp, welcher häufig zu seiner Ausprägung das Vorhandensein zusätzlicher nicht-genetischer Umweltfaktoren voraussetzt.

35

Der heute existierende Ansatz solche DNA Regionen zu identifizieren, welche bei der Krankheitsentwicklung eine Rolle spielen, basiert auf der Betrachtung polymorpher, genetischer Marker und statistischen Analysen, welche nach einer Korrelation zwischen bestimmten Markerallelen und einem Krankheitsphänotyp suchen. Die meisten dieser Verfahren basieren auf dem Nachweis von Regionen "identical-by-descent (IBD)" d.h. Regionen, welche bei zwei verwandten



Individuen mit dem gleichen Krankheitsphänotyp durch Abstammung identisch sind und demnach eine gewisse, größer als zufällige, Wahrscheinlichkeit haben, an der Expression des betrachteten Phänotyps beteiligt zu sein (Tran LD, Elston RC, Keats BJB & Wilson AF. Sib-pair linkage program (SIBPAL) in S.A.G.E Statistical Analysis of Genetic Epidemiology, Release 2.2 (Louisiana State University, New Orleans, 1994. Haseman JK & Elston RC. The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet* 2, 3-19 (1972)). Dieses Verfahren durchgeführt an vielen oft hunderten von Familien kann zum Nachweis von chromosomalen an der Krankheit beteiligten genetischen Loci führen.

Der Nachweis solcher Regionen erfolgt normalerweise durch den Gebrauch sogenannter Mikrosatellitenmarker. Diese hoch-polymorphen DNA Abschnitte, welche durch eine Anzahl repetitiver, kurzer Sequenzabschnitte (z.B. CA, CAAT etc.) gekennzeichnet sind, lassen sich leicht mit Hilfe der Polymerase Chain Reaktion (PCR) und anschließenden Separationsverfahren über Polyacrylamidgele nachweisen (C. Dib et al. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* (1996) 380:152-154). Dies erfolgt heute meist über ein semi-automatisiertes Verfahren in welchem die Oligonukleotide zur PCR Amplifikation mit Fluorophoren markiert werden und dann von einem automatischen Sequenziergerät nachgewiesen werden können. Vielfach wird ein solches Experiment genomweit durchgeführt. Hierfür werden Marker in geeigneten Abständen (normalerweise etwa 10 centi Morgan (cM)) über das gesamte Genom verteilt selektioniert, durch PCR amplifiziert und die individuellen Allele wie oben beschrieben nachgewiesen (Davies et al. A genome-wide search for human type 1 diabetes susceptibility genes. *Nature*. (1994) 371:130-6. , Hager et al. A genome-wide screen for human obesity genes reveals a major susceptibility locus on chromosome 10, *Nature Genet* 20: 304-308 (1998)).

Der Nachteil dieses Verfahrens liegt in der Komplexität der betrachteten Phänotypen. Um eine ausreichende statistische Aussagekraft zu erhalten müssen deswegen bei diesem Verfahren hunderte, oft auch tausende von Individuen getestet werden. Der damit verbundene Aufwand ist enorm (Hodge S.E. Linkage analysis versus association analysis: distinguishing between two models that explain disease-marker associations. *Am J Hum Genet*. (1993) 53 :367-84, Lander E. & Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genet* 11: 241-247 (1995)). Darüberhinaus ist der materielle Aufwand und die damit verbundenen Kosten hoch. Außerdem liefert dieses Verfahren auch im besten Fall nur eine relativ geringe Auflösung der Region, in welcher sich das Gen (bzw. die Gene) befinden, welche bei der Krankheitsentwicklung eine Rolle spielen. Oft ist diese Region größer als 10 cM



(dies entspricht im Mittel etwa 10 Millionen Nukleobasen). Das eigentliche Auffinden des Gens (der Gene) erfordert daher nachfolgende, langwierige Klonierungs-verfahren. Diese beinhalten das Erstellen einer physikalischen Karte der Region (z.B. mit Hilfe artifizierter, bakterieller Chromosomen, BAC's) und viele aufwendige und teure Sequenzierschritte.

5

Die oben beschriebenen Schwierigkeiten der klassischen Nachweismethode haben in neuerer Zeit deswegen das Interesse an neuen Verfahren geweckt. Besonderes Augenmerk fällt dabei auf eine Methode, welche als "Genomische Mismatch Analyse" (Genomic Mismatch Scan, GMS) bezeichnet wird (Nelson S. et al. *Genet* 4: 11-18 (1993)). Dieses Verfahren zeichnet sich durch einige Besonderheiten aus, welche die oben beschriebenen Nachteile der klassischen Methodik ausräumen. Wie die klassische Methode beruht dieses Verfahren auf dem Umstand, daß chromosomale Regionen, welche durch eine genetische Veränderung zur Krankheitsprädisposition beitragen, bei zwei verwandten Individuen mit dem gleichen Krankheitsphänotyp durch Abstammung identisch (IBD) sein sollten. Bei zwei verwandten Individuen sind gewisse chromosomale Regionen IBD, da sie von den gleichen genetischen Vorfahren stammen. Mit dem Verwandtschaftsgrad nimmt auch der Anteil dieser Regionen ab, d.h. der Grad an Homologie der DNA zwischen den Individuen verringert sich mit ihrem verwandtschaftlichen Abstand. Dabei sollte aber eine Region, welche bei der Ausprägung eines Krankheitsphänotyps eine Rolle spielt auch über abnehmende Verwandtschaftsverhältnisse IBD erhalten bleiben. Im Mittel enthält die menschliche DNA alle 100-1000 Basen eine polymorphe Stelle (Cooper D.N. et al. (1985) An estimate of unique DNA sequence heterozygosity in the human genome. *Hum Genet* 69:201-205). Nimmt man nun die DNA von zwei verwandten Individuen des gleichen Phänotyps (z.B. Verwandte dritten, vierten oder fünften Grades) und hybridisiert diese miteinander dann sollten nur wenige Regionen ihrer genomischen DNA's 100% IBD sein (Zwei Kusinen haben im Mittel z.B. 1/16 ihrer DNA IBD). Unter den homologen Fragmenten sollten sich aber solche Loci befinden, welche an der Ausprägung des entsprechenden Phänotyps beteiligt sind.

10

20

25

30

35

Technisch wird diese Methode durchgeführt, indem man genomische DNA von zwei verwandten Individuen durch Restriktionsenzyme in kleinere Fragmente schneidet, mischt, durch Erhitzen denaturiert und dann re-naturieren läßt. Eine der beiden DNA's wird dabei vor der Hybridisierung enzymatisch vollständig methyliert. Das Ergebnis dieser Reaktion sind vollständig- und nicht-methylierte Homohybride, sowie Heterohybride mit einem methylierten und einem



unmethylierten DNA Strang, welche den re-naturierten DNA's beider Individuen entsprechen. Die vollständig- bzw. nicht-methylierten, nicht informativen Homohybride können durch geeignete spezifische Enzyme (z.B. DpnI und MboI) abgebaut werden. Die verbleibenden Heterohybride bilden ein Gemisch aus

5 vollständig homologen Fragmenten und solchen, welche unterschiedliche Basensequenzen aufweisen und daher 'mismatches' beinhalten. Diese Mismatches können durch geeignete Enzyme, wie z.B. das E. coli Reparaturenzym MutS, erkannt und die mismatch enthaltende DNA eliminiert werden, so daß zuletzt nur die

10 vollständig homologen DNA Fragmente, welche IBD sind erhalten bleiben. Diese können dann auf verschiedene Weise analysiert und charakterisiert werden. Ein Enzymsystem, welches sich als besonders geeignet für dieses Verfahren erwiesen hat ist das MutHLS mismatch Reparationssystem. Dies besteht aus drei interagierenden Proteinen, MutS, welches Mismatches erkennt und an der Stelle des

Mismatches an die DNA bindet, MutL, welches an der DNA entlangfährt bis es eine spezifische Erkennungssequenz (GATC) findet und MutH, welches hinter der Erkennungssequenz einen Einzelstrangbruch erzeugt (Pang et al. The mutH, mutL

mutS and uvrD genes of Salmonella typhimurium LT2. Cold Spring Harbor Symposia on quantitative Biology (1984) 49:597-602, Modrich et al. DNA mismatch correction. *Ann Rev Biochem* (1987) 56:435-466, Smith J and Modrich

20 P. (1996) Mutation detection with MutH, MutL and MutS mismatch repair proteins. *Proc. Natl. Acad. Sci. USA* 93: 4374-4379. (Nelson S. et al. (1993) Genomic mismatch scanning, a new approach to genetic linkage scanning *Nature Genet* 4: 11-18. Cheung V.G. Nelson F.S. (1998) Genomic mismatch scanning identifies human genomic DNA shared identical by descent *Genomics* 47: 1-6).

25 Der Vorteil von GMS gegenüber der klassischen Markeranalyse besteht auf zwei Ebenen. Zum einen ist der Nachweis ein direkter bei dem keine aufwendigen statistischen Analysen benötigt werden. Im Idealfall reichen zwei verwandte Individuen gleichen Phänotyps einer Familie aus, um die Prozedur erfolgreich durchzuführen. Damit ist der logistische und finanzielle Aufwand zur Sammlung

30 von Patientenmaterial erheblich verringert. Zum anderen werden bei der Durchführung dieses Verfahrens durch die Restriktion handliche, definierte DNA Fragmente erzeugt (z.B. werden bei einem Restriktionsenzym mit einer häufigen 4-Basen Erkennungssequenz im Mittel Fragmente von 3000 Basen erzeugt). Diese Fragmente lassen sich leicht analysieren und charakterisieren.

35 .

Aufgabenstellung

5 Derzeit lassen sich mit der GMS Methode genetische Loci für multifaktorielle
Phänotypen nicht nachweisen. Dies liegt zum einen an der komplexen Natur dieser
Phänotypen, welche kein klares den Mendelschen Gesetzen folgendes
Vererbungsmuster zeigen zum anderen an der Komplexität der notwendigen
molekularen Reaktionen selbst, wie z.B. der Komplexität der menschlichen DNA
10 und der Effizienz des eingesetzten Enzymsystems. So ist z.B. die notwendige
Separation der Mismatch enthaltenden Heterohybride an das Vorhandensein der
spezifischen Erkennungssequenz für MutL (GATC) gebunden. Dies bedeutet, daß
Fragmente, welche mismatches enthalten aber nicht diese Erkennungssequenz, nicht
eliminiert werden und erhöhen die falsch-positiv Rate des Experiments. Die
Komplexität der Krankheit erschwert ausserdem die Selektion der wirklich am
Phänotyp beteiligten Fragmente von solchen, welche IBD aber ohne phänotypische
Bedeutung sind.

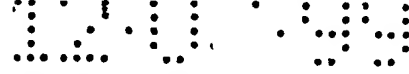
20 Das der vorliegenden Erfindung zugrundeliegende und zu lösende technische
Problem war demnach, ein effizientes und kostengünstiges Verfahren zur
Identifikation von IBD Loci für komplexe phänotypische Merkmale bereitzustellen,
welche diese Limitationen der GMS Anwendung ausräumen.

25 Dieses technische Problem wird durch die Bereitstellung der in den
Patentansprüchen gekennzeichneten Ausführungsformen gelöst.

Beschreibung

30 Das Verfahren dient zur Identifikation chromosomaler Regionen IBD zwischen
Individuen mit dem gleichen komplexen Phänotyp (z.B. einer komplexen
Erkrankung). Für das Verfahren sind die folgenden Schritte durchzuführen. a)
Restriktion der DNA von mindestens zwei, verwandten Individuen, b) Ligation
spezifischer Adaptormoleküle an die Restriktionsfragmente, c) Amplifikation der
generierten Fragmente durch PCR, d) Methylierung der DNA einer der an der
35 Reaktion der beteiligten Individuen, e) Hybridisierung der Fragmente von jeweils
zwei Individuen und Abbau von renaturierten Homohybriden durch
methylabhängige bzw. methylinaktivierte Restriktionsendonukleasen (z.B. Dpn1
und Mbo 1) und f) durch eine spezifische Enzymreaktion werden " Mismatches "
identifiziert und die entsprechenden Fragmente eliminiert.

- Das erfindungsgemäße Verfahren erlaubt die spezifische Amplifikation der zu untersuchenden DNA's durch eine PCR Reaktion um so genügend Ausgangsmaterial zu erhalten. Dies wird dem erfindungsmäßigen Verfahren entsprechend durch Addition spezifischer Adaptormoleküle (Oligonukleotide bekannter Sequenz) nach Restriktion der DNA erreicht. Diese Adaptormoleküle haben den Vorteil, daß hierdurch nicht nur die Restriktionsfragmente gezielt amplifiziert werden können, sondern, daß sich die Adaptorensequenzen so wählen lassen, daß zusätzliche Informationen in die Fragmente eingebracht werden können, wie die Erkennungssequenzen für die bei der Eliminierung der Homohybride zu verwendenden Restriktionsendonukleasen. Weiter können die Adaptoren z.B. die Erkennungssequenz für das MutHLS System integrieren, um so einen vollständigen Abbau der Mismatch enthaltenden Fragmente zu erreichen. Andere sequenzspezifische Elemente der Adaptoren sind z.B. Sequenzen für weitere Restriktionsendonukleasen, welche die spätere Direktklonierung der Fragmente in geeignete Vektoren (z.B. Plasmide) zulassen. Auch können solche Adaptoren mit Molekülen wie z.B. Biotin markiert werden, welche eine Immobilisierung der resultierenden PCR Fragmente an einer Festphase zulassen.
- In einer weiteren bevorzugten Ausführungsart des erfindungsgemäßen Verfahrens wird die direkte Amplifikation mit kurzen Oligonukleotiden unterschiedlicher Sequenz-zusammensetzung (Random Priming) erreicht. Dieses Verfahren läßt mit einigen sequenzgegebenen Ausnahmen die gleichen Manipulationen zu wie die Adaptoren Variante. Hier kann jedoch auf eine vorherige Restriktion der DNA verzichtet werden.

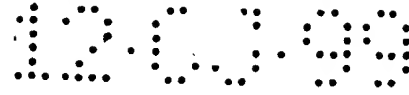


Eine weitere bevorzugte Variante des erfindungsgemäßen Verfahren kombiniert in vorteilhafter Weise zytogenetische und molekulargenetische Verfahren, welche es erlauben die Komplexität des GMS Verfahrens so herabzusetzen, daß dieses Verfahren routinemäßig angewandt werden kann um chromosomale Regionen IBD, auch solche welche an der Entwicklung komplexer Phänotypen beteiligt sind, zu identifizieren.

Eine weitere Möglichkeit die Komplexität der durchzuführenden Hybridisierungsreaktionen herabzusetzen kann durch die Vorselektion von Chromosomen der zu untersuchenden Individuen erreicht werden. Dies kann manuell unter einem Phasenkontrastmikroskop geschehen, indem von einem Objekträger etwa 20-40 G-Band gefärbte Chromosomen mittels einer Glaspipette entfernt und gesammelt werden. In einer weiteren Variante kann die Abtrennung der Chromosomen automatisch über einen Zellsorter, welcher auch die Trennung der Chromosomen erlaubt, erfolgen. Durch diese Trennung wird die nachfolgende GMS Reaktion wesentlich vereinfacht und spezifischer. Diese Chromosomenpräparation ist das Ausgangsmaterial für die oben beschriebenen Verfahrensschritte.

In einer bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens werden nach erfolgter PCR Amplifikation der Gesamt-DNA oder der Chromosomenpräparation aus dem Gemisch der amplifizierten DNA Fragmente repetitive Elemente weitgehend eliminiert. Diese Eliminierung kann z.B. über die Schmelzcharakteristik der DNA Fragmente erreicht werden. Hierzu wird die DNA durch Hitze denaturiert und unter stringenten Bedingungen (Puffersystem, Temperatur) re-naturiert. Repetitive Elemente zeigen eine größere v_{max} bei der Renaturierung als nicht-repetitive Elemente. Die Zeit, in der die meisten repetitiven aber nicht die uni-Sequenzfragmente renaturiert haben, läßt sich experimentell ermitteln. Die verbleibenden einzelsträngigen Uni-sequenzen können stabilisiert werden (z.B. durch niedrige Temperatur). Diese Einzelstränge können dann von den doppelsträngigen repetitiven Elementen z.B. durch Bindung an Hydroxylapatit oder andere Einzelstrangspezifische Moleküle von der doppelsträngigen DNA getrennt werden (Sedat J.W., Kelly R.B. & Sinsheimer R.L. (1967) Fractionation of nucleic acid on benzoylated-naphthoylated DEAE cellulose. *J Mol Biol* 26:537-540).

In einer weiteren bevorzugten Variante des erfindungsgemäßen Verfahrens können repetitive Elemente durch Absättigung mit kurzen Sonden der gängigen repetitiven Sequenzen (Alu, Di-, Tri, Tetra nukleotidsequenzen) aus dem Gemisch entfernt



werden. Dazu können diese Sonden mit Molekülen markiert werden, welche die Bindung an eine feste Matrix erlauben (z.B. Biotin).

5 In einer bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens wird zu dem die Heterohybride enthaltenen Gemisch nach erfolgter Amplifikation und Abbau der Homohybride das Protein MutS zugegeben, um an die Mismatches zu binden. Der Protein/DNA Komplex wird dann bei dieser bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens direkt an eine proteinbindende Oberfläche (z.B. PVDF-Membran, reversed phase Säule)
10 gebunden und so von der nicht-komplexierten DNA getrennt.

In einer weiteren bevorzugten Variante des Verfahrens wird der MutS Protein/DNA Komplex an spezifische Antikörper gebunden und durch eine Immunopräzipitation von der freien DNA getrennt. Die Antikörper werden in einer bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens an eine Matrix innerhalb einer Säule gebunden. Bei der Anwendung des erfindungsgemäßen Verfahrens wird die an MutS gebundene DNA dann über diese Säule durch die Antikörperbindung von den nicht MutS gebundenen, vollständig homologen DNA Fragmenten getrennt.

20

Beispiele

Beispiel 1: Direkter Nachweis von IBD Fragmenten aus genomischer DNA

25 Genomische DNA von drei verwandten Individuen, von denen zwei sich durch den gleichen Phänotyp und das dritte durch die Abwesenheit dieses Phänotyps auszeichnen, wird nach Standardmethoden (z.B. Phenol-Chloroform Extraktion) extrahiert. Die DNA's werden getrennt mit einer relativ häufig schneidenden Restriktionsendonuklease (z.B. Pst1) geschnitten. Zu diesen
30 Restriktionsfragmenten werden dann Adaptoren hinzupipettiert, welche für den entstandenen Restriktionsüberhang komplementär sind und an die Restriktionsfragmente ligiert. Die Sequenz der Adaptoren ist dabei so gewählt, daß sie die folgenden Elemente enthalten: a) Erkennungssequenzen für methylabhängige bzw. Methylinaktivierte Endonukleasen (z.B. Dpn1 und Nde1), b)
35 Erkennungssequenz für MutH und c) die Erkennungssequenz für eine Restriktionsendonuklease (z.B. Kpn1), welche die Klonierung in einen Plasmidvektor des Typs BlueScript zulassen. Nach erfolgter Hybridisierung und Ligation der Adaptoren an die Restriktionsfragmente werden die Fragmente durch

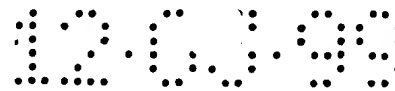


Zugabe adaptorspezifischer, komplementärer Oligonukleotide durch PCR amplifiziert.

Die Amplifikationsprodukte eines des an der Reaktion beteiligten Individuums mit dem betrachteten Phänotyp wird dann mittels Dam Methylase an den Adenin Basen methyliert. Danach werden jeweils die PCR Produkte dieses Individuums mit denen der anderen beiden Individuen zusammenpipettiert. Durch Hitzedenaturierung und anschließende Renaturierung der DNA Fragmente in einem speziellen Reaktionspuffer werden Heteroduplexes aus der DNA der zwei verschiedenen Individuen gebildet, welche auf einem der beiden DNA Stränge eine Methylierung aufweisen (Casna et al. (1986) genomic analysis II, isolation of high molecular weight heteroduplex DNA following methylase protection and formamide PERT hybridization *Nucleic Acids Res.* 14: 7285-7303). Daneben entstehen aus der Hybridisation der DNA's der beiden Individuen untereinander vollständig methylierte und nicht-methylierte Homohybride. Zu diesen Hybrid DNA's wird eine methylabhängige Endonuklease (z.B. Dpn1) sowie eine methylinaktivierte Endonuklease (z.B. Nde1) hinzugegeben. Durch das Einbringen der Erkennungssequenzen für die beiden Enzyme in die Adaptoren wird ein vollständiger Verdau der Homohybride gewährleistet. Die geschnittenen Homohybridmoleküle werden dann durch eine Exonuklease (z.B. Exo3) weiter abgebaut und durch Bindung an eine Einzelstrangspezifische Matrix (z.B. BNDC) aus dem Gemisch entfernt.

Zu den verbliebenen Hybrid DNA Molekülen werden MutS, MutL und MutH hinzupipettiert. MutS bindet an etwaige Mismatches innerhalb der DNA Fragmente und propagiert die Bindung von MutL und MutH was zum Schneiden der DNA an der spezifischen Erkennungssequenz GATC führt. Da diese Sequenz als Teil des Adaptormoleküls in die DNA eingeführt wurde, wird gewährleistet, daß alle Mismatch enthaltenden DNA Fragmente geschnitten werden. Wie nach Restriktion der Homohybride werden die Schnitte durch eine Exonuklease erweitert und über ein einzelstrang-bindendes System (z.B. BNDC) aus dem Gemisch entfernt. Durch dieses Verfahren wurde eine Anreicherung von absolut homologen IBD Fragmenten erreicht.

Die Analyse der verbliebenen Fragmente wird über ihre direkte Klonierung in einen Vektor (z.B. BlueScript) erreicht. Die nach der bakteriellen Amplifikation erhaltenen Klone können z.B. der Direktsequenzierung zugeführt werden oder durch Hybridisierung mit einem geordnetem cDNA Array auf kodierende Sequenzen hin untersucht werden.



Beispiel 2: Direkter Nachweis von IBD Fragmenten nach Chromosomenspezifischer Amplifikation

5 Ausgehend von Chromosomenpräparationen G-band gefärbter Mitosechromosomen von drei verwandten Individuen, von denen zwei sich durch den gleichen Phänotyp und das dritte durch die Abwesenheit dieses Phänotyps auszeichnen, werden von jedem Individuum 20-40 Exemplare eines jeden Chromosoms mittels einer Glaspipette isoliert. Für jedes der 23 Chromosomenexemplare wird ein getrenntes Gefäß (z.B. Eppendorf- Reaktionsgefäß) bereitgestellt, in welches die 20-40

10 isolierten Chromosomen überführt werden.

Die so nach Chromosomen geordneten DNA's werden für jedes Individuum getrennt mit einer relativ häufig schneidenden Restriktionsendonuklease (z.B. PstI) geschnitten. Zu diesen Restriktionsfragmenten werden dann Adaptoren hinzupipettiert, welche für den entstandenen Restriktionsüberhang komplementär sind und an die Restriktionsfragmente ligiert. Die Sequenz der Adaptoren ist dabei so gewählt, daß sie die folgenden Elemente enthalten: a) Erkennungssequenzen für DpnI und NdeI, b) Erkennungssequenz für MspI und c) die Erkennungssequenz für eine Restriktionsendonuklease (z.B. KpnI), welche die Klonierung in einen Plasmidvektor des Typs BlueScript zulassen. Nach erfolgter Ligation der Adaptoren

20 werden die Fragmente durch Zugabe adaptorspezifischer, komplementärer Oligonukleotide durch PCR amplifiziert.

Die chromosomenspezifischen Amplifikationsprodukte eines der beiden Individuen mit dem betrachteten Phänotyp wird dann mittels Dam Methylase an den Adenin Basen methyliert. Danach werden jeweils die PCR Produkte dieses Individuums

25 immer getrennt nach Chromosomen mit denen der anderen beiden Individuen zusammenpipettiert d.h. bei der Hybridisierung von drei Individuen werden 46 (2x23 Chromosomen) unabhängige Hybridisierungsreaktionen vorgenommen. Durch Hitzedenaturierung und anschließender Renaturierung der DNA Fragmente in einem speziellen Reaktionspuffer werden Heteroduplexe aus der DNA der zwei

30 verschiedenen Individuen gebildet, welche auf einem der beiden DNA Stränge eine Methylierung aufweisen (Casna et al. (1986) Genomic analysis II, isolation of high molecular weight heteroduplex DNA following methylase protection and formamide PERT hybridization *Nucleic Acids Res.* 14: 7285-7303). Daneben entstehen aus der Hybridisation der DNA's der beiden Individuen untereinander vollständig

35 methylierte und nicht-methylierte Homohybride. Zu diesen Hybrid DNA's wird die methylinabhängige Endonuklease DpnI, sowie die methylinaktivierte Endonuklease NdeI hinzugegeben. Durch das Einbringen der Erkennungssequenzen für die beiden Enzyme in die Adaptoren wird so ein vollständiger Verdau der Homohybride gewährleistet. Die geschnittenen Homohybridmoleküle werden dann durch eine



Exonuklease (z.B. Exo3) weiter abgebaut und durch Bindung an eine einzelstrangspezifische Matrix (z.B. BNDC) aus dem Gemisch entfernt.

Zu den verbliebenen Hybrid DNA Molekülen werden MutS, MutL und MutH hinzupipettiert. MutS bindet an etwaige Mismatches innerhalb der DNA Fragmente und propagiert die Bindung von MutL und MutH, was zum Schneiden der DNA an der spezifischen Erkennungssequenz GATC führt. Da diese Sequenz als Teil des Adaptormoleküls in die DNA eingeführt wurde, wird gewährleistet, daß alle Mismatch enthaltenden DNA Fragmente geschnitten werden. Wie nach Restriktion der Homohybride werden die Schnitte durch eine Exonuklease erweitert und über BNDC aus dem Gemisch entfernt. Durch dieses Verfahren wird eine Anreicherung von absolut homologen IBD Fragmenten erreicht.

Die Analyse der verbliebenen Fragmente wird über ihre direkte Klonierung in einen Vektor (z.B. BlueScript) erreicht. Die nach der bakteriellen Amplifikation erhaltenen Klone können z.B. der Direktsequenzierung zugeführt werden oder durch Hybridisierung mit einem geordnetem cDNA Array auf kodierende Sequenzen hin untersucht werden.

Beispiel 3: Direkte Isolierung Mismatch enthaltender Heteroduplexes über proteinbindende Säulen.

Genomische DNA von drei verwandten Individuen, von denen zwei sich durch den gleichen Phänotyp und das dritte durch die Abwesenheit dieses Phänotyps auszeichnen, wird nach Standardmethoden (z.B. Phenol-Chloroform Extraktion) extrahiert. Die DNA's werden getrennt mit einer relativ häufig schneidenden Restriktionsendonuklease (z.B. Pst1) geschnitten. Zu diesen Restriktionsfragmenten werden dann Adaptoren hinzupipettiert, welche für den entstandenen Restriktionsüberhang komplementär sind und an die Restriktionsfragmente ligiert. Die Sequenz der Adaptoren ist dabei so gewählt, daß sie die folgenden Elemente enthalten: a) Erkennungssequenzen für Dpn1 und Nde1, b) Erkennungssequenz für MutH und c) die Erkennungssequenz für eine Restriktionsendonuklease (z.B. Kpn1), welche die Klonierung in einen Plasmidvektor (z.B. BlueScript) zulassen. Nach erfolgter Ligation der Adaptoren werden die Fragmente durch Zugabe adaptorspezifischer, komplementärer Oligonukleotide durch PCR amplifiziert.

Die Amplifikationsprodukte eines der beiden Individuen mit dem betrachteten Phänotyp wird dann mittels Dam Methylase an den Adenin Basen methyliert. Danach werden jeweils die PCR Produkte dieses Individuums mit denen der anderen beiden Individuen zusammenpipettiert. Durch Hitzedenaturierung und anschließende

Renaturierung der DNA Fragmente in einem speziellen Reaktionspuffer werden Heteroduplexes aus der DNA der zwei verschiedenen Individuen gebildet, welche auf einem der beiden DNA Stränge eine Methylierung aufweisen (Casna et al. (1986) genomic analysis II, isolation of high molecular weight heteroduplex DNA following methylase protection and formamide PERT hybridization *Nucleic Acids Res.* 14: 7285-7303). Daneben entstehen aus der Hybridisation der DNA's der beiden Individuen untereinander vollständig methylierte und nicht-methylierte Homohybride. Zu diesen Hybrid DNA's wird die methylabhängige Endonuklease Dpn1, sowie die methylinaktivierte Endonuklease Nde1 hinzugegeben. Durch das Einbringen der Erkennungssequenzen für die beiden Enzyme in die Adaptoren wird so ein vollständiger Verdau der Homohybride gewährleistet. Die geschnittenen Homohybridmoleküle werden dann durch eine Exonuklease (z.B. Exo3) weiter abgebaut und durch Bindung an eine einzelstrangspezifische Matrix (z.B. BNDC) aus dem Gemisch entfernt.

Zu den verbliebenen Heterohybriden wird MutS hinzupipettiert, welches an " mismatch " enthaltende Fragmente bindet. Über eine Säule, in welcher der DNA/Protein Komplex nicht aber die freie DNA an eine Matrix gebunden wird, wird die " mismatch " enthaltende DNA, gebunden an das mismatch erkennende Protein, aus dem Gemisch entfernt.

Die Analyse der verbliebenen Fragmente wird über ihre direkte Klonierung in einen Vektor (z.B. BlueScript) erreicht. Die nach der bakteriellen Amplifikation erhaltenen Klone können z.B. der Direktsequenzierung zugeführt werden oder durch Hybridisierung mit einem geordnetem cDNA oder Mikrosatelliten Array auf kodierende Sequenzen hin untersucht werden.

Beispiel 4: Anwendung der GMS Methode bei der Zuchtoptimierung von landwirtschaftlich genutzten Haustieren

Ein wichtiges Gebiet der modernen Tierzucht ist die Optimierung bestimmter, quantitativer Merkmale von landwirtschaftlichen Nutztieren (z.B. Rindern, Schweinen, Schafen). Auch hier kann das vorgestellte erfindungsgemäße Verfahren vorteilhaft angewandt werden. Das folgende Beispiel zeigt eine mögliche Anwendung bei der Qualitätsoptimierung für Milch in der Käseverarbeitung.

Ein wichtiges quantitatives Merkmal bei der Auswahl von Milch für die Verarbeitung zu Käse ist der Caseingehalt. Je höher dieser liegt desto größer ist die Käseausbeute. Bei der Anwendung von GMS wird genomische DNA von mindestens zwei verwandten Tieren, ausgezeichnet durch einen

überdurchschnittlichen Caseingehalt der Milch, nach Standardmethoden (z.B. Phenol-Chloroform) extrahiert und mit einer häufig schneidenden Restriktionsendonuklease (z.B. PstI) verdaut. Zu diesen Restriktionsfragmenten werden dann Adaptoren hinzupipettiert, welche für den entstandenen Restriktionsüberhang komplementär sind und mittels Hybridisierung und Zugabe einer Ligase an die Restriktionsfragmente ligiert werden. Die Sequenz der Adaptoren ist dabei so gewählt, daß sie mindestens die folgenden Sequenzelemente enthalten: a) Erkennungssequenzen für methylabhängige bzw. Methylinaktivierte Endonukleasen (z.B. DpnI und NdeI) und b) Erkennungssequenz für den MutLH Komplex. Nach erfolgter Ligation der Adaptoren werden die DNA Fragmente durch Zugabe adaptorspezifischer, komplementärer Oligonukleotide durch PCR amplifiziert.

Die Amplifikationsprodukte der DNA eines der beiden Tiere, welches hohe caseinhaltige Milch gibt, wird dann mittels Dam Methylase an den Adenin Basen methyliert. Danach werden jeweils die PCR Produkte dieses Tieres mit denen des anderen Tieres zusammenpipettiert. Durch Hitzedenaturierung und anschließende Renaturierung der DNA Fragmente in einem speziellen Reaktionspuffer werden Heteroduplexes aus der DNA der zwei verschiedenen Individuen gebildet, welche auf einem der beiden DNA Stränge eine Methylierung aufweisen (Casna et al. (1986) genomic analysis II, isolation of high molecular weight heteroduplex DNA following methylase protection and formamide PERT hybridization *Nucleic Acids Res.* 14: 7285-7303). Daneben entstehen aus der Hybridisation der DNA's der beiden Individuen untereinander vollständig methylierte und nicht-methylierte Homohybride. Zu diesen Hybrid DNA's wird eine methylabhängige Endonuklease (z.B. DpnI), sowie eine methylinaktivierte Endonuklease (z.B. NdeI) hinzugegeben. Durch das Einbringen der Erkennungssequenzen für die beiden Enzyme in die Adaptoren wird ein vollständiger Verdau der Homohybride gewährleistet. Die geschnittenen Homohybridmoleküle werden dann durch eine Exonuklease (z.B. Exo3) weiter abgebaut und durch Bindung an eine einzelstrangspezifische Matrix (z.B. BNDC) aus dem Gemisch entfernt. Zu den verbliebenen Hybrid DNA Molekülen werden MutS, MutL und MutH hinzupipettiert. MutS bindet an etwaige Mismatches innerhalb der DNA Fragmente und propagiert die Bindung von MutL und MutH, was zum Schneiden der DNA an der spezifischen Erkennungssequenz GATC führt. Da diese Sequenz als Teil des Adaptormoleküls in die DNA eingeführt wurde, wird gewährleistet, daß alle Mismatch enthaltenden DNA Fragmente geschnitten werden. Wie nach Restriktion der Homohybride werden die Schnitte durch eine Exonuklease erweitert und über BNDC aus dem Gemisch entfernt. Durch dieses Verfahren wird eine Anreicherung von absolut homologen IBD Fragmenten erreicht unter denen sich auch die Loci

befinden, welche für die Caseinproduktion verantwortlich sind. Diese können durch Bindung z.B. an ein cDNA Array identifiziert werden.

Beispiel 5: Identifizierung Krankheitsrelevanter IBD Fragmente nach GMS mittels Mikrosatellitenanalyse auf einem DNA array

5

10

Genomische DNA von drei verwandten Individuen, von denen zwei sich durch den gleichen Krankheitsphänotyp und das dritte durch die Abwesenheit dieses Phänotyps auszeichnen, wird nach Standardmethoden (z.B. Phenol-Chloroform Extraktion) extrahiert. Die DNA's werden getrennt mit einer relativ häufig schneidenden Restriktionsendonuklease (z.B. Pst1) geschnitten. Zu diesen Restriktionsfragmenten werden dann Adaptoren hinzupipettiert, welche für den entstandenen Restriktionsüberhang komplementär sind und an die Restriktionsfragmente ligiert. Die Sequenz der Adaptoren ist dabei so gewählt, daß sie mindestens die folgenden Elemente enthalten: a) Erkennungssequenzen für methylabhängige bzw. Methylinaktivierte Endonukleasen (z.B. Dpn1 und Nde1), b) Erkennungssequenz für den MutLH Komplex und c) eine chemische Aktivität (z.B. Radioaktivität, Fluorophor), welche den Nachweis der Fragmente erlaubt. Nach erfolgter Hybridisierung und Ligation der Adaptoren an die Restriktionsfragmente werden die Fragmente durch Zugabe adaptorspezifischer, komplementärer Oligonukleotide durch PCR amplifiziert.

20

25

30

35 .

Die Amplifikationsprodukte eines der Individuen mit dem betrachteten Krankheitsphänotyp wird dann mittels Dam Methylase an den Adenin Basen methyliert. Danach werden jeweils getrennt die PCR Produkte dieses Individuums mit denen der anderen beiden Individuen zusammenpipettiert. Durch Hitzedenaturierung und anschließende Renaturierung der DNA Fragmente in einem speziellen Reaktionspuffer werden Heteroduplexes aus der DNA der zwei verschiedenen Individuen gebildet, welche auf einem der beiden DNA Stränge eine Methylierung aufweisen (Casna et al. (1986) genomic analysis II, isolation of high molecular weight heteroduplex DNA following methylase protection and formamide PERT hybridization *Nucleic Acids Res.* 14: 7285-7303). Daneben entstehen aus der Hybridisation der DNA's der beiden Individuen untereinander vollständig methylierte und nicht-methylierte Homohybride. Zu diesen Hybrid DNA's wird eine methylabhängige Endonuklease (z.B. Dpn1), sowie eine methylinaktivierte Endonuklease (z.B. Nde1) hinzugegeben. Durch das Einbringen der Erkennungssequenzen für die beiden Enzyme in die Adaptoren wird so ein vollständiger Verdau der Homohybride gewährleistet. Die geschnittenen Homohybridmoleküle werden dann durch eine Exonuklease (z.B. Exo3) weiter



abgebaut und durch Bindung an eine einzelstrangspezifische Matrix (z.B. BNDC) aus dem Gemisch entfernt.

5 Zu den verbliebenen Hybrid DNA Molekülen werden MutS, MutL und MutH hinzupipettiert. MutS bindet an etwaige Mismatches innerhalb der DNA Fragmente und propagiert die Bindung von MutL und MutH was zum Schneiden der DNA an der spezifischen Erkennungssequenz GATC führt. Da diese Sequenz als Teil des Adaptormoleküls in die DNA eingeführt wurde, wird gewährleistet, daß alle Mismatch enthaltenden DNA Fragmente geschnitten werden. Wie nach Restriktion der Homohybride werden die Schnitte durch eine Exonuklease erweitert und über
10 BNDC aus dem Gemisch entfernt. Durch dieses Verfahren wurde eine Anreicherung von absolut homologen IBD Fragmenten erreicht.

Auf einer festen Oberfläche (z.B. einem kommerziell erhältlichen Glaschip oder in einer Mikrotiterplatte mittels kovalenter Bindung) werden spezifische Oligonukleotide (oder zur Erhöhung der Spezifität Peptidnukleinsäuren (PNA's)) für eine definierte Anzahl von Mikrosatelliten gebunden. Gegen dieses DNA-array werden die verbliebenen Produkte aus der GMS Reaktion hybridisiert. Nach entsprechenden Waschschritten sind nach diesem Vorgang solche GMS Produkte an den chip gebunden, welche homologe Bereiche zu den mikrosatellitenmarkerspezifischen Oligonukleotiden aufweisen. Diese können durch
20 die in die Adaptorsequenz eingebrachte chemische Aktivität nachgewiesen werden. Der Vergleich der Hybridisierungsexperimente zwischen dem krank-kranken und krank-gesunden Paar, wird dabei zur Identifikation krankheitsrelevanter Fragmente genutzt.

25

Ansprüche

30 1. Verfahren zur Identifikation von chromosomalen Regionen "identical by descent" zwischen Individuen mit gleichen quantitativen oder qualitativen phänotypischen Merkmalen, ausgezeichnet dadurch, daß folgende Verfahrensschritte ausgeführt werden:

35 a) die DNA von mindestens zwei, verwandten Individuen durch Restriktionsendonukleasen geschnitten werden, b) spezifische Adaptormoleküle, ausgezeichnet durch die Einführung von Sequenzen, welche den Erkennungssequenzen einer methyl-abhängigen Restriktionsendonuklease (z.B. DpnI) und eines methylinaktivierten Restriktionsenzym, sowie der Erkennungssequenz für z.B. MutL entsprechen, an die Restriktionsfragmente ligiert werden, c)

5 die generierten Fragmente durch eine PCR amplifiziert werden, d) die DNA eines der an der Reaktion beteiligten Individuen methyliert wird, e) Heteroduplexes aus den Amplifikationsprodukten zweier Individuen gebildet werden, f) renaturierte Homohybriden durch methylabhängige bzw. methylinaktivierte Restriktionsendonukleasen abgebaut werden, g) durch eine spezifische Reaktion ein "mismatch" erkennendes Protein an die DNA gebunden und dadurch entfernt wird h) und die übriggebliebene DNA-Fragmente, die für die untersuchten Phänotypen charakteristischen IBD-DNAs enthält.

10 2. Verfahren nach Anspruch 1. dadurch ausgezeichnet, daß es sich bei der in a) geschnittenen DNA um sortierte Chromosomen handelt, welche in jeweils getrennten Reaktionsgefäßen weiterbehandelt werden.

3. Verfahren nach Anspruch 1. dadurch ausgezeichnet, daß in die in Verfahrensschritt b) eingesetzten Adaptormoleküle eine Restriktionsschnittstelle zum direktklonieren der amplifizierten Produkte enthalten.

20 4. Verfahren nach Anspruch 1. dadurch ausgezeichnet, daß das "mismatch" bindende Enzym in g) MutS ist.

5. Verfahren nach den Ansprüchen 1. - 4. dadurch ausgezeichnet, daß nach der Bindung von MutS an die DNA diese durch Zugabe von MutL und MutH geschnitten wird.

25 6. Verfahren nach Anspruch 1. dadurch ausgezeichnet, daß anstelle der in b) beschriebenen Adaptormoleküle ein "random priming" mit kurzen Oligonukleotiden durchgeführt wird, welche die spezifischen Erkennungssequenzen tragen und eine Amplifikation ohne vorherige Ligation erlauben.

30 7. Verfahren nach Anspruch 1. bei dem die in b) hergestellten Adaptormoleküle mit einer chemischen Funktion versehen werden, welche die Bindung an eine Oberfläche zuläßt.

35 8. Verfahren nach Anspruch 1. dadurch ausgezeichnet, daß nach der nach Anspruch 4. erfolgten Bindung von MutS an die mismatch enthaltenden DNA's diese Protein/DNA Hybride über eine Festphase, welche die spezifische Bindung des Proteins erlaubt, aus der Reaktion entfernt werden.

9. Kit für das Verfahren nach den Ansprüchen 1. - 8., bestehend aus der Kontroll-DNA von 3 verschiedenen Individuen zur Positivkontrolle, Restriktionsendonukleasen um einen spezifischen Verdau durchzuführen, spezifische Adaptoren mit den in 1. Ausgeführten Sequenzspezifikationen, um die IBD Fragmente zu identifizieren.
10. Kit für das Verfahren nach den Ansprüchen 1. - 9. einen DNA-chip zur Identifikation der IBD Fragmente enthaltend.
11. Kit für das Verfahren nach Anspruch 10., dadurch ausgezeichnet, dass die Oligonukleotidsequenzen auf dem DNA-chip den Primersequenzen bekannter Mikrosatellitenmarker entsprechen.